

## I le vocabulaire des statistiques.

### définition 1 :

L'ensemble sur lequel on travaille en statistique est appelé **population**.

Si cet ensemble est trop vaste, on en restreint l'étude à une partie appelée **échantillon**.

Un élément de cet ensemble est appelé **individu**.

### définition 2 :

La particularité commune que l'on étudie est appelée **caractère**.

L'**effectif** d'une « valeur » d'un caractère est le nombre d'individus ayant cette valeur.

Les valeurs prises par le caractère sont aussi appelées les **modalités**.

(1) si celle-ci s'exprime par un nombre, il s'agit d'un **caractère quantitatif**. (dans ce cas, le nombre se note en général  $x_i$ )

Si les valeurs du nombre exprimé en (1) sont isolées, il s'agit d'un caractère discret.

Par contre, si ces valeurs sont prises dans tout un intervalle de  $\mathbb{R}$ , il s'agit d'un **caractère continu**.

### exemples :

Le nombre de frère et sœur d'un élève de seconde 10 est un caractère **quantitatif discret**.

il peut prendre les valeurs 0, 1, 2, 3, 4 ....

La taille des élèves de seconde 10 est un caractère **quantitatif continu**. il peut prendre toutes les valeurs entre 1,50 m et 1,95 m.

(2) si la particularité étudiée ne s'exprime pas par un nombre, il s'agit d'un caractère **qualitatif**.

### exemple :

Dans une population, être marié(e) est un caractère qualitatif à deux valeurs : oui ou non.

### définition 3 :

Une série statistique est l'ensemble des résultats d'une étude : valeurs du caractère et effectifs correspondants.

On représente souvent une série statistique sous forme d'un tableau.

### définition 4 :

Le nombre d'individus ( $n_i$ ) d'une modalité est appelé effectif.

Le nombre total d'individus ( $N$ ) de la population est appelé effectif total.

Le rapport  $f_i = \frac{n_i}{N}$  est appelé **fréquence**.

### remarques :

$f_i$  est un nombre toujours compris entre 0 et 1.

Souvent, les nombres  $f_i$  s'expriment par un pourcentage.

La somme des nombres  $f_i$  est toujours égale à 1.

## II caractéristiques d'une série.

### définition 1 :

La différence des valeurs extrêmes du caractère s'appelle l'**étendue**.

### exemple :

Si on étudie des notes, c'est la différence entre la note la plus haute et la note la plus basse.

remarque : on ne parle d'étendue que dans le cas d'un caractère quantitatif.

définition 2 :

le mode pour un caractère discret, noté  $M_o$ , est la valeur qui correspond au plus grand effectif. dans le cas d'un caractère continu, on parle de classe modale.

Dans le cas de classes de même amplitude, la classe modale est une classe qui correspond au plus fort effectif.

définition 3 : la moyenne

dans le cas d'une série discrète.

la moyenne est le nombre noté  $\bar{x}$

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N} \quad \text{ou} \quad \bar{x} = \frac{f_1x_1 + f_2x_2 + \dots + f_px_p}{N}$$

si les valeurs sont regroupées en classe, on calcule la moyenne en choisissant comme valeurs du caractère les centres des classes et comme effectifs, les effectifs des classes.

définition 4 :

la médiane d'une série, dont les valeurs du caractères sont discrètes et rangées dans l'ordre croissant, est la valeur partageant la population en deux groupes de même effectif.

exemples :

1)

note : $x_i$	5	8	12	15	18
nombre d'élèves : $n_i$	7	5	14	3	2

le mode est la note 12 car l'effectif 14 est le plus grand.

la médiane est la note du 16ème élève car il y a 31 élèves. c'est donc 12.

la moyenne :

$$\bar{x} = \frac{7 \cdot 5 + 5 \cdot 8 + 14 \cdot 12 + 3 \cdot 15 + 2 \cdot 18}{7 + 5 + 14 + 3 + 2} = \frac{324}{31} \approx 10,45.$$

l'étendue est :  $18 - 5 = 13$ .

2)

taille (en cm)	[1,5 ; 1,6[	[1,6 ; 1,7[	[1,7 ; 1,8[	[1,8 ; 1,9[	[1,9 ; 2[
effectif	5	16	9	4	1

la classe [1,6 ; 1,7[ est la **classe modale** car elle a le plus grand effectif. son centre, 1,65 est le mode de la série.

la moyenne :

$$\bar{x} = \frac{5 \cdot 1,55 + 16 \cdot 1,65 + 9 \cdot 1,75 + 4 \cdot 1,85 + 1 \cdot 1,95}{5 + 16 + 9 + 4 + 1} = \frac{59,25}{35} \approx 1,69$$